



Catalogue no. 2280

Home Language and Knowledge of Languages

1996 Census Technical Reports



Census



Statistique
Canada

Statistics
Canada

Canada

How to Reach Us

Data in Many Forms...

Statistics Canada disseminates data in a variety of forms. In addition to publications, both standard and special tabulations are offered. Data are available on CD, diskette, computer print-out, microfiche and microfilm, and magnetic tape. Maps and other geographic reference materials are available for some types of data. Direct on line access to aggregated information is possible through CANSIM, Statistics Canada's machine-readable data base and retrieval system.

How to Obtain More Information

Inquiries about products and services should be directed to your local Statistics Canada Reference Centre:

Halifax	(902 426-5331)	Regina	(306 780-5405)
Montréal	(514 283-5725)	Edmonton	(403 495-3027)
Ottawa	(613 951-8116)	Calgary	(403 292-6717)
Toronto	(416 973-6586)	Vancouver	(604 666-3691)
Winnipeg	(204 983-4020)		

Toll-free access is provided in all provinces and territories, for users who reside outside the local dialing area of any of the regional reference centres.

Newfoundland, Labrador, Nova Scotia, New Brunswick and Prince Edward Island	1-800-565-7192
Quebec	1-800-361-2831
Ontario	1-800-263-1136
Saskatchewan	1-800-667-7164
Manitoba	1-800-661-7828
Southern Alberta	1-800-882-5616
Alberta and the Northwest Territories	1-800-563-7828
British Columbia and the Yukon Territory	1-800-663-1551
Telecommunications Device for the hearing impaired	1-800-363-7629
Toll Free Order Only Line (Canada and United States)	1-800-267-6677

How to Order Publications

Statistics Canada publications may be purchased from local authorized agents and other community bookstores, through the nearest Statistics Canada Regional Reference Centre, or by mail order to Marketing Division, Sales and Service, Statistics Canada, Ottawa K1A 0T6. Contact the nearest Statistics Canada Regional Reference Centre for more information on electronic products and services. See the order form at the end of this product for further details.

(1-613-951-7277)
Facsimile Number (1-613-951-1548)
Toronto Credit Card Only (1-416-973-8018)

Standards of Service to the Public

To maintain quality service to the public, Statistics Canada follows established standards covering statistical products and services, delivery of statistical information, cost-recovered services and service to respondents. To obtain a copy of these service standards, please contact your nearest Statistics Canada Regional Reference Centre.



Statistics Canada

1991 Census Technical Reports

Home Language and Knowledge of Languages

Reference Products series

Published by authority of the Minister responsible for Statistics Canada

©Minister of Industry, Science and Technology, 1994

All rights reserved. No part of this document may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without prior written permission from Licence Services, Marketing Division, Statistics Canada, Ottawa, Ontario, Canada K1A 0T6.

February 1995

Price: Canada: \$20.00
United States: US\$24.00
Other Countries: US\$28.00

Catalogue 92-336E

ISBN 0-660-14269-4

Ottawa

La version française de cette publication est disponible sur demande (92-336F)

Note of Appreciation

Canada owes the success of its statistical system to a long-standing cooperation involving Statistics Canada, the population of Canada, its businesses and governments. Accurate and timely statistical information could not be produced without their continued cooperation and goodwill.

Canadian Cataloguing in Publication Data

Main entry under title:

Home Language and Knowledge of Languages

(1991 census technical reports)

(Reference products series)

Issued also in French under title: Langue parlée
à la maison et connaissance des langues.

ISBN 0-660-14269-4

CS92-336E

1. Canada -- Languages -- Statistics.
 2. Canada -- Census, 1991 -- Methodology.
 3. Canada -- Census, 1991 -- Evaluation.
- I. Statistics Canada. II. Series.

HA741.5.1991 H65 1995 306.4'0971021

C94-988090-6

How to Cite This Document

Statistics Canada. Home Language and Knowledge of Languages.

1991 Census Technical Reports; Reference Products Series.

Ottawa: Minister of Industry, Science and Technology, 1995.

Catalogue number 92-336E.

Preface

Through time, the Census of Canada has become the primary source of information about Canadians and how they live. Decisions based on this information affect the social and economic affairs of all Canadians.

Statistics Canada, as the professional agency in charge of producing this information, has the responsibility for informing users of data quality. The agency must describe the concepts and methodology used in collecting and processing the data, as well as any other features that may affect their use or interpretation.

In order to describe the quality of the 1991 Census data, Statistics Canada has prepared the following publications: a census **Dictionary**, which provides concise and easy to understand textual and graphical information pertaining to census concepts; a **Handbook**, which provides an overview of how the census is conducted; and a series of **Technical Reports**, which present in greater detail, information on the quality of data for specific characteristics, such as industry, as covered in this report.

Information on data quality is important for users. It allows them to assess the usefulness of census data for their purposes as well as the risks involved in basing conclusions or decisions on these data. The 1991 Census was a large and complex undertaking and, while considerable effort was taken to ensure high standards throughout all collection and processing operations, the resulting data are inevitably subject to a certain degree of error.

Information on data quality is also important to Statistics Canada. It is an integral part in the development and maintenance of pertinent and reliable statistical programs.

This publication is a major contribution to achieving these goals. It has been prepared by Brian Harrison, with the technical assistance of Bob McCrea. Preliminary work on the report was completed by Raj Chawla, with technical assistance from Luc St-Amour, while on an assignment with the Demolinguistics Division. The support of staff from two divisions in Statistics Canada was also essential for completing this report: Census Operations and Social Survey Methods.

Finally, I would like to express my appreciation to the millions of Canadians who completed their questionnaires on June 4, 1991, as well as to those who assisted Statistics Canada in planning and conducting the census.

Ivan P. Fellegi
Chief Statistician of Canada

Table of Contents

	Page	
List of Tables	ii	
I. Introduction	1	
II. Concepts and Definitions	2	
III. Data Collection and Coverage	6	
IV. Data Assimilation	14	
V. Edit and Imputation	17	
VI. Data Evaluation	27	
VII. Historical Comparability	29	
VIII. Products and Services	35	
IX. Conclusion	38	
Regional Reference Centres	39	

List of Tables

Tables	Page
3.1 Estimated Net Undercoverage for Mother Tongue, Home Language and Knowledge of Official Languages, 1991	8
3.2 Non-response Rates for Language Variables, Canada, Provinces and Territories, 1991	12
3.3 Number and Percentage of Multiple Responses for Mother Tongue and Home Language, Canada, Provinces and Territories, 1991	13
4.1 Write-in Responses by Method of Coding for Mother Tongue, Home Language and Knowledge of Non-official Languages, Canada, 1991	16
5.1 Corrections of Anomalies and Non-response, Mother Tongue and Home Language, Canada, 1991 ..	19
5.2 Correction of Anomalies and Non-response, Knowledge of Official Languages, Canada, 1991	21
5.3 Correction of Anomalies and Non-response, Knowledge of Non-official Languages, Canada, 1991 ..	22
5.4 Home Language Before and After Edit and Imputation, Canada, 1991	24
5.5 Knowledge of Official Languages Before and After Edit and Imputation, Canada, 1991	25
5.6 Knowledge of Non-official Languages Before and After Edit and Imputation, Canada, 1991	25
7.1 Mother Tongue, Home Language, Knowledge of Official Languages and Knowledge of Non-official Languages for Permanent and Non-permanent Residents, Canada, 1991	30
7.2 Principal Languages Affected by the Inclusion of Non-permanent Residents, Canada, 1991	32
7.3 Multiple Response for Mother Tongue and Home Language, Canada, 1981, 1986, 1991	33

I. Introduction

Every five years a census of population is carried out in Canada. The national Census of Population is a major project conducted by Statistics Canada to collect, verify and publish data. The national Census provides the most comprehensive database on the characteristics of Canadians, their families and their households. The information ranges from age and sex of individuals to their ethnic origin, education, occupation, labour force activity, industry, sources of income, their family and household characteristics. The census is an invaluable source of information that is useful to the various levels of government, to businesses, associations, educational institutions, interest groups, and to the general public. The data can be used in government planning of social and economic programs, assessment of the need for educational and health facilities, and planning by private enterprise.

Information is obtained through a series of questions established after detailed consultation and testing. It is collected by trained enumerators, checked for inconsistencies and errors and forwarded to Revenue Canada Taxation (RCT) regional centres for data entry. The final results are placed on a computer database at Statistics Canada. Data are analyzed, published and disseminated in various forms.

In a massive project such as the census, however, the results are never perfect. Although considerable effort has been made to maintain high standards of quality, errors inevitably occur at various stages of the collection and processing operations. Users must be aware of the nature and scope of any errors that the census data may contain, as well as the risks involved in basing conclusions or decisions on these data.

The **1991 Census Technical Reports** have been designed to inform data users of the potential problems or intricacies of the data. The reports inform users of the conceptual framework and definitions used in the data collection, any unusual circumstances which may influence the data, likely principal sources of error and, where possible, the size of the error.

This product is a specialized analytical tool. It complements and co-ordinates other reference products and assists the more sophisticated user to understand variable details and methodological information on coverage, sampling and weighting.

II. Concepts and Definitions

The definitions of census terms, variables and concepts are presented here as they appear in the **1991 Census Dictionary** (Catalogue No. 92-301E). Users should refer to the 1991 Census Dictionary for full definitions and additional remarks related to any concepts and definitions not found in this chapter.

Census Variables

While the 1991 Census consisted of fifty-three (53) questions, the **1991 Census Dictionary** (Catalogue No. 92-301E or D) lists well over 200 variables. Obviously, there is not a one-to-one correspondence between questions and variables. Several variables can be produced from one question only, while other variables are derived using responses from several questions.

By the same token, some census variables closely resemble information as it was gathered on the questionnaires while others are very different. For example, sex has two answer categories, male and female: the categories on the questionnaire correspond exactly to those on the database. **Sex** is therefore called a **direct** variable.

Derived variables have undergone transformations. For example "date of birth" was asked on the questionnaire but **age** is the database variable. Sometimes the link between collected information and the database variable is not so straightforward. For example, several questions are used to identify the **unemployed**, yet the word "**unemployed**" does not appear anywhere on the questionnaire, and its definition is not intuitively obvious.

Census variables are grouped into these categories:

- counts and demographic data;
- ethnic origin and immigration data;
- language;
- aboriginal status;
- schooling;
- religion;
- labour force;
- income;
- families;
- housing and households;
- institutions and other collectives;
- disability.

The potential for creating new census variables is virtually limitless. Some variables can be very conventional and **direct**, while other variables can be **derived** and tailored to user needs.

Universes

A "**universe**" in the census refers to what is counted in a tabulation. The possibilities are:

- population (i.e., persons);
- families;

- households;
- dwellings.

The **Population Universe** includes variables that provide information about individuals. It covers a wide variety of characteristics such as demographic, ethno-cultural, language, mobility, schooling, income and labour force. A complete list of these variables can be found in the Table of Contents of the *1991 Census Dictionary* (Catalogue No. 92-301E). Some variables within this universe are collected for the entire population of Canada; others are collected for a sample of the population only.

The primary objective of the Census is to provide accurate coverage of the entire population of Canada at various geographic levels. The 1991 Census provided counts for:

- all Canadian citizens and landed immigrants with a residence in Canada;
- Canadian citizens and landed immigrants who are abroad, either on a military base or attached to a diplomatic mission;
- Canadian citizens and landed immigrants at sea or in port aboard merchant vessels under Canadian registry;
- **non-permanent residents** (persons who hold student or employee authorizations, Minister's permits or who are refugee claimants).

"Families" are groups within a household. Within the Family Universe two general categories are identified: **census families** and **economic families**.

- A **Census Family** refers to a now-married couple (with or without never-married sons and/or daughters of either or both spouses), a couple living common-law (again with or without never-married sons and/or daughters of either or both partners), or a lone parent of any marital status, with at least one never-married son or daughter living in the same dwelling.
- An **Economic Family** refers to a group of two or more persons who live in the same dwelling and are related to each other by blood, marriage, common-law or adoption. For example, a brother and a sister living together; or a mother and her separated daughter; would constitute an economic family, but not a census family.

The **Household Universe** is composed of subuniverses and variables which pertain to a person or a group of persons (other than temporary or foreign residents) who occupy a dwelling. Examples of household universes are private households, collective households, households outside Canada.

The **Dwelling Universe** is composed of subuniverses (collective and private) and variables pertaining to characteristics of dwellings in Canada. Dwellings are distinct from households. Dwelling characteristics refer to physical attributes of a set of living quarters, whereas household characteristics pertain to the person or group of persons (other than foreign and/or temporary residents) who occupy a dwelling.

Census Geography

Statistics Canada uses a very accurate and detailed geographic structure that makes it possible to obtain information for many different geographical units, known as geographic areas. Data from the 1991 Census is available for numerous standard geographic areas, as well as non-standard or user-defined areas.

Census Boundaries

In order to take a census for a country as large as Canada, smaller geographic boundaries must be established to facilitate enumeration. The basic boundaries are the provinces (PROV's), the federal electoral districts (FED's) and, finally, a smaller unit called the enumeration area (EA).

Standard Geographic Areas

Census data are disseminated for a number of standard geographic areas. These areas are of two (2) types: legislative/administrative and statistical.

- (a) Legislative/administrative areas are defined, with a few exceptions, by Canadian federal and provincial statutes. These include:

Geographic Area	Total Number
• Provinces and Territories;	12
• Federal Electoral Districts (FEDs);	295
• Census Divisions (CDs);	290
• Census Subdivisions (CSDs);	6,006
• Subprovincial Regions (SPRs).	68

- (b) Statistical areas are defined by Statistics Canada as part of the spatial frame used to collect and disseminate census data. These include:

Geographic Area	Total Number
• Agricultural Regions;	76
• Census Consolidated Subdivisions (CCSs);	2,630
• Census Metropolitan Areas (CMAs);	25
• Census Agglomerations (CAs);	115
• Primary Census Metropolitan Areas (PCMAs);	12
• Primary Census Agglomerations (PCAs);	21
• Census Tracts (CTs);	4,068
• Provincial Census Tracts (PCTs);	1,815
• Urban Areas (UAs)/Rural Areas;	893
• CMA/CA Parts	N/A
• CMA/CA Components;	N/A
• Enumeration Areas (EAs).	45,995

Other geographic units of quasi-standard nature are **unincorporated place (UP)**, **township**, **range and meridian** and **postal code**.

User-defined Areas

Census data can also be produced for areas other than the standard geographic areas, that is for user-defined areas. These are of two (2) types: aggregation of standard geographic areas and custom **query areas**.

An in depth look at terms related to the geography of the 1991 Census are defined in the geography section of the **1991 Census Dictionary**. They describe, more extensively, concepts related to geographic areas and census cartography.

Demolinguistic Concepts

The United Nations recommends that three types of information be collected on language:

- (a) mother tongue, defined as the language usually spoken in the individual's home in his early childhood;
- (b) usual language, defined as the language currently spoken by the individual in his present home; and
- (c) ability to speak one or more designated languages.

Canadian census practices have followed these recommendations fairly closely. In fact, Canada is one of the only countries in which census data are collected on all three aspects. Data on mother tongue, home language and knowledge of languages have been collected for over one hundred languages, in addition to the two official languages.

Home Language

The home language refers to the language spoken most often at home by the individual at the time of the census. It should be noted that another language, or indeed more than one language, may be spoken by the individual at home, but the home language is the one spoken most often.

Knowledge of Languages

The 1991 Census included two questions on knowledge of languages. The first dealt with knowledge of official languages, i.e. English and French. The second, which appeared for the first time, dealt with knowledge of non-official languages.

Knowledge of Official Languages

The variable refers to the ability of the respondent to conduct a conversation in English only, in French only, in both English and French, or in none of the official languages of Canada.

The data on knowledge of official languages are based on respondent assessment of his or her ability to speak the two official languages.

Knowledge of Non-official Languages

This variable refers to the ability to conduct a conversation in languages other than English or French. The non-official language data are based on the respondent assessment of his or her ability to speak non-official languages.

This question was asked for the first time in the 1991 Census.

Mother Tongue

Refers to the first language learned at home in childhood and still understood by the individual at the time of the census.

III. Data Collection and Coverage

For the 1991 Census, information was collected from more than 11 million dwellings both in Canada and abroad. The data collection process consists of the drop-off and retrieval of approximately 11,500,000 questionnaires. These questionnaires are then edited to ensure they have been properly completed by all Canadians across the country. This phase employed approximately forty thousand (40,000) people in a variety of tasks from mapping to post-censal activities.

Two **collection methods** were used for the 1991 Census: **self-enumeration** and **canvasser enumeration**. In self-enumeration areas, a questionnaire (Form 2A or Form 2B) was dropped off at each household before Census Day (June 4th). A member of the household was to complete the questionnaire on Census Day. Questionnaires were mailed back in pre-addressed envelopes. In 1991, less than 2% of households were enumerated by canvassers: Census Representatives completed a long form questionnaire (Form 2D) for these households by interview. This method was used to enumerate each household in remote or northern areas and on Indian reserves where irregular mail service makes mail-back impractical. Some of the remote areas were enumerated as early as March 1991. Data was collected on *every Canadian citizen, landed immigrant and non-permanent resident alive at midnight between June 3rd and June 4th 1991*.

The two main types of accepted **enumeration approaches** used by census takers are the "**de jure**" approach and "**de facto**" approach. The "**de jure**" approach assigns the person to the dwelling in which he/she usually resides. The "**de facto**" approach assigns the person to the dwelling in which he/she is staying at the time of enumeration. In Canada the "**modified de jure**" approach is used. It allows a subsequent matching process to determine whether or not a person enumerated in one dwelling as temporary was also identified as a usual resident in his/her stated usual place of residence. This approach recognizes, and compensates for the potential failure of a straight "**de jure**" approach to enumerate persons away from their usual place of residence on census night.

In 1991, a number of initiatives were taken to improve coverage. These included:

- Using paid advertising to inform Canadians on when and how "to count themselves in";
- Creating an address register from other sources of information and using this list to check if any dwellings were missed;
- Establishing special procedures to count homeless people through soup kitchens;
- Establishing special procedures to count the population on Indian reserves;
- Respondent-friendly questionnaire;
- Public Communications Program and a multilingual Telephone Assistance Service.
- Edit and follow-up by CR for non-response and missing information
- Quality checks of the CR's assigned by the CC and the Quality Control Technician

The primary components, documents and geographical unit used for data collection and coverage are:

- **Visitation Record (VR) (Form 1)**

A document used by Census Representatives to list the household numbers and the number of persons per household including temporary residents in all enumeration areas. The VR lists every private and collective dwelling (occupied and unoccupied), as well as every agricultural holding in an enumeration area. The VR also provides control totals to help ensure that all dwellings and persons were enumerated.

- **Census of Population Questionnaires**

Seven questionnaires have been designed for the collection of data. The 1991 Census Questionnaires include:

Collective Dwelling Record	(Form 1A);
Short Questionnaire	(Form 2A);
Long Questionnaire	(Form 2B);
Overseas Population Form	(Form 2C);
Long Canvasser Questionnaire	(Form 2D);
Individual Census Questionnaire	(Form 3);
Soup Kitchen Questionnaire	(Form 3B).

- **Enumeration Area (EA)**

An EA is an area canvassed by a Census Representative. It is the smallest geographical unit for which census data are available. The number of dwellings vary from 375 (maximum) in large urban areas to 125 (minimum) in rural areas.

Please refer to the **Long Questionnaire (Form 2B)** for questions asked in the 1991 Census of Canada.

Address Register (AR)

In 1991, to help increase coverage, an Address Register (AR) was implemented in urban areas with populations of fifty thousand (50,000) and over. The Address Register is a list compiled by merging several administrative data files from a variety of sources. The Census Representative was to compare the addresses in the Visitation Record when the questionnaires were delivered with the addresses in the list taken from the Address Register. If an address appearing in the AR did not match any address in the Visitation Record, the Census Representative returned to the sector to locate the dwelling at the missed address. Conversely, any address entered by the Census Representative that was not in the AR was added to the AR after the census.

Special Coverage Studies

Since one hundred percent (100%) coverage is virtually impossible in such a large survey, a number of checks are performed on the collection of data. These studies measure the extent of coverage errors that occur when dwellings or individuals are missed, incorrectly included, or double-counted. Some examples of these checks are the **Vacancy Check**, **Temporary Residents Study**, **Reverse Record Check**, and the **Overcoverage Study**. These studies will be discussed in more detail in the following chapters on Data Assimilation (Chapter IV) and Data Evaluation (Chapter VI), and in the technical report on **Coverage** (Catalogue No. 92-341).

The accuracy of the counts for the census are assessed by using the **Estimated Net Undercoverage**. This measure indicates the effect of the two types of coverage errors (i.e. undercoverage minus overcoverage). While the net undercoverage rate for the entire population was 2.9%, there was considerable variation for subgroups of respondents. New immigrants and young adults, for example, are often missed by the census.

Table 3.1 shows net undercoverage estimates for three language variables – Mother Tongue, Home Language and Knowledge of Official Languages. Estimates for the Knowledge of Non-official Languages variable are not available. It is apparent that the undercoverage rates are considerably higher for languages other than English or French for both the mother tongue and home language variables. The mother tongue results show that net undercoverage rates are very similar for the French and English mother tongue groups (approximately 2.5%), but are almost double for the non-official languages (4.9%). While Italian and German language groups have low rates of net undercoverage for both mother tongue and home language, the Chinese group and the “other” languages have higher rates of net

undercoverage (approximately 6%). This is partly attributable to the greater tendency of recent immigrants, many of whom have Chinese or "other" mother tongue and home language, to be missed by the census. The net undercoverage rate for home language is similar for those with English home language and those with French home language.

Table 3.1 Estimated Net Undercoverage for Mother Tongue, Home Language and Knowledge of Official Languages, 1991

Characteristics	Net number of persons missed		Population net undercoverage rate	
	Estimated number	Standard error	Estimate rate (%)	Standard error (%)
Mother Tongue				
Total	807,254	37,971	2.87	0.13
English	438,395	35,520	2.49	0.20
French	172,927	16,258	2.48	0.23
Non-official languages	206,854	15,653	4.85	0.35
Italian	5,861	3,915	1.13	0.74
German	10,518	4,145	2.16	0.83
Chinese	31,143	5,777	5.95	1.03
Other	159,027	12,564	5.77	0.43
Home Language				
Total	802,353	43,148	2.89	0.15
English	505,798	39,351	2.64	0.20
French	156,955	18,990	2.40	0.28
Non-official languages	148,617	14,220	5.72	0.51
Italian	5,241	3,652	1.79	1.21
German	3,736	2,089	2.70	1.46
Chinese	28,632	5,476	6.24	1.11
Other	110,943	11,138	6.48	0.60
Knowledge of Official Languages				
Total	802,353	43,148	2.89	0.15
English only	531,507	35,518	2.85	0.18
French only	88,167	15,720	2.10	0.36
English and French	136,298	15,588	3.01	0.33
Neither English nor French	46,380	7,175	10.92	1.49

Source: 1991 Census Technical Report, *Coverage*, Catalogue No. 92-341E

- Note:
1. For home language and mother tongue, multiple responses were accepted, and therefore the estimates of net undercoverage for individual categories do not necessarily add to the totals, e.g., a response of "English and French" would be counted once as English and once as French.
 2. The grand totals for home language and knowledge of official languages differ slightly from the total for mother tongue, since these two grand totals exclude institutional residents.

The results for the knowledge of official languages variable show that those who speak neither official language have a very high rate of net undercoverage (approximately 11%). This may be attributable to the recency of arrival of this population.

The aboriginal languages are affected by particular undercoverage difficulties. On some Indian reserves and Indian settlements, enumeration was not permitted or was interrupted before it could be completed. Data for 1991 are therefore not available for the incompletely enumerated Indian reserves and Indian settlements and are not included in tabulations. These Indian reserves and Indian settlements (a total of 78) were more likely to be inhabited by people who speak an aboriginal language, and the estimates for mother tongue, home language and knowledge of official languages are affected accordingly. Counts for aboriginal languages must be interpreted with caution, especially in geographic areas containing the incompletely enumerated Indian reserves and Indian settlements (they are listed in all 1991 Census Nation Series and Dimension Series publications).

Questions and Instructions

The 2A questionnaire, which was completed by approximately 80% of Canadian households, included only one language question (mother tongue). However, the 2B questionnaire contained four language questions dealing with knowledge of official languages, knowledge of non-official languages, home language and mother tongue.

The first question on knowledge of official languages asked respondents about their ability to speak English and French.

LANGUAGE

7. Can this person speak English or French well enough to conduct a conversation?

Mark one circle only.

07.

- 01 ☐ English only
02 ☐ French only
03 ☐ Both English and French
04 ☐ Neither English nor French

The following instructions, designed to assist respondents in ambiguous circumstances, appeared in the census guide which accompanied the 2B questionnaire.

Do not report English or French learned at school unless a conversation of some length on various topics can be carried on in that language.

For a child who has not yet learned a language, report the language used **at home** if it is either English or French. If both languages are used **often**, mark the circle labelled **Both English and French**. If neither language is used **often**, mark **Neither English nor French**.

If deaf or mute, report English, French or both provided that one or both of these languages is understood.

Note: Research has shown that many respondents do not refer to the guide when completing the questionnaire.

Knowledge of Non-official Languages

This question, similar to the question on official languages, appeared directly after it on the census questionnaire.

8. What language(s), other than English or French, can this person speak well enough to conduct a conversation?

05 ☐ None
OR
Specify other language(s)

06 _____
07 _____
08 _____

The following instructions, pertaining to this question, appeared in the census guide.

If no languages other than English or French are spoken, mark the circle labelled **None**.

Do not report a language learned at school unless a conversation of some length on various topics can be carried on in that language.

For a child who has not yet learned a language, report a language **other than English or French** if used often at home.

If deaf or mute, report a sign language if used.

If a person speaks an **Indian** language (from **India**), do not report Indian but rather the specific language such as Hindi, Urdu or Punjabi.

If a person speaks an **Amerindian** (North American Indian) language, report the specific Amerindian language such as Cree or Ojibway.

Home Language

The question on home language, the language spoken most often at home, appeared after the question on knowledge of non-official languages and before the question on mother tongue.

9. What language does this person speak most often at home?

09 ☐ English

10 ☐ French

Other - Specify
11 _____

The following instructions, appearing in the guide, provided clarification for respondents in special circumstances.

Report the language each person speaks most often at home. If you live alone, report the language in which you feel most comfortable. If two languages are used **equally often**, report both.

For a child who has not yet learned a language, report the language used most often at home. If two languages are used **equally often**, report **both**.

If deaf or mute, report the language used most often at home to communicate.

If a person speaks an **Indian** language (from **India**), do not report Indian as the language spoken at home but rather the specific language such as Hindi, Urdu or Punjabi.

If a person speaks an **Amerindian** (North American Indian) language, report the specific Amerindian language such as Cree or Ojibway.

Mother Tongue

The mother tongue question was the last language question on the 2B questionnaire.

10. What is the language that this person first learned at home in childhood and still understands?

If this person no longer understands the first language learned, indicate the second language learned.

12 ☐ English

13 ☐ French

14 ☐ Other - Specify

The following instructions, pertaining to mother tongue, were included in the census guide.

Report the first language learned **at home** before starting school. If this language is no longer understood, report the second language learned.

If more than one language is understood, report the language first learned at home. If two languages were learned at the same time, report the language spoken **most often as a child at home**. If both languages were used **equally often**, report **both**.

For a child who has not yet learned a language, report the first language this child will learn **at home**. If two languages will be learned at the same time, report the one which will be used most often to speak to this child. If both languages are used **equally often**, report **both**.

If a person speaks an **Indian** language (from **India**), do not report Indian but rather the specific language such as Hindi, Urdu or Punjabi.

If a person speaks an **Amerindian** (North American Indian) language, report the specific Amerindian language such as Cree or Ojibway.

Non-response for the Language Questions

Non-response often results when respondents cannot provide the census information, having been away throughout the census period. More frequently, the questionnaire is returned but information is missing for some questions or individuals. Considerable effort is devoted to ensure as complete a response as possible. A toll-free telephone number is provided to respondents in the event that they require assistance. The census representatives (CRs) edit the questionnaires and follow up on missing information. The questions on mother tongue and knowledge of official languages were both designated for mandatory follow-up by the CR.

Table 3.2 shows the non-response rates for the four language questions. The non-response rate was very low (0.7%) for the knowledge of official languages question, somewhat higher (1.2%) for the questions on mother tongue and home language, and much higher (3.8%) for the question on knowledge of non-official languages. The latter question was often skipped by respondents who apparently did not think that the question applied to them, and did not check the box "none" to show that they did not know any non-official languages. The low non-response rate for the question on knowledge of official languages may be attributable to its placement on the 2B questionnaire (it was the first language question).

Table 3.2 Non-response Rates for Language Variables, Canada, Provinces and Territories, 1991

	Mother tongue	Home language	Knowledge of official languages	Knowledge of non-official languages
Canada	1.2	1.2	0.7	3.8
Newfoundland	0.8	0.8	0.4	2.6
Prince Edward Island	0.8	0.9	0.4	3.5
Nova Scotia	0.8	0.9	0.5	2.9
New Brunswick	0.8	0.9	0.4	3.4
Quebec	1.1	1.0	0.6	4.3
Ontario	1.4	1.3	0.8	3.8
Manitoba	1.5	1.3	0.9	3.5
Saskatchewan	1.0	0.9	0.6	2.7
Alberta	1.1	1.0	0.6	3.2
British Columbia	1.4	1.3	0.8	3.8
Yukon	5.0	4.6	4.2	6.9
Northwest Territories	2.5	2.2	1.4	3.5

Note: 1. Based on weighted 20% sample data.

2. The rates refer to the proportion of records for which no response was provided. An invalid response is counted as a response for these calculations.

Response rates were generally lowest in the eastern provinces and highest in the Yukon. Among the provinces, the rates for mother tongue ranged from 0.8% in the four eastern provinces to 1.5% in Manitoba; the rates for home language from 0.8% in Newfoundland to 1.3% in Ontario, Manitoba and British Columbia; the rates for knowledge of official languages from 0.4% in Newfoundland, Prince Edward Island and New Brunswick to 0.9% in Manitoba; and the rates for knowledge of non-official languages from 2.6% in Newfoundland to 4.3% in Quebec.

Multiple Response

The instructions for the mother tongue question in the census guide indicated that two languages should be reported if they were learned at the same time in early childhood and both were used equally often. The instructions for the home language question indicated that two languages should be reported if they were used equally often.

In total, 330,250 multiple responses were reported for mother tongue and 487,730 for home language. The tendency to provide multiple responses was lowest in the three eastern provinces, which are characterized by considerable linguistic homogeneity, somewhat higher in New Brunswick and markedly higher in all other provinces and territories.

Table 3.3 Number and Percentage of Multiple Responses for Mother Tongue and Home Language, Canada, Provinces and Territories, 1991

	Mother tongue		Home language	
	Number	%	Number	%
Canada	330,250	1.2	487,730	1.8
Newfoundland	590	0.1	835	0.1
Prince Edward Island	365	0.3	360	0.3
Nova Scotia	3,135	0.4	3,245	0.4
New Brunswick	4,750	0.7	6,020	0.8
Quebec	75,035	1.1	127,160	1.9
Ontario	141,875	1.4	212,355	2.1
Manitoba	20,000	1.9	24,195	2.2
Saskatchewan	11,855	1.2	11,895	1.2
Alberta	32,805	1.3	40,650	1.6
British Columbia	38,345	1.2	58,415	1.8
Yukon	320	1.2	260	0.9
Northwest Territories	1,175	2.0	2,330	4.1

Note: Based on weighted 20% sample data.

IV. Data Assimilation

Data assimilation is the processing phase during which data from the Census questionnaires are edited, coded and captured. The process includes the transformation of the questionnaire responses into machine-readable form.

The four main components of data assimilation are:

- Regional Office Processing
- Direct Data Entry
- Head Office Processing
- Automated Coding

Regional Office Processing (ROP)

At this stage, ROP staff will ensure that information appearing on the questionnaires is suitable for key entry into the computer. This operation employs approximately 2,000 people, and is conducted in Revenue Canada - Taxation (RCT) regional processing centres in St. John's, Jonquière, Shawinigan, Sudbury, Winnipeg and Surrey. In Ottawa it is conducted in the Statistics Canada head office. For the 1991 Census, the operation took place during the period between July and November of 1991.

ROP operations consisted of the following:

(a) Receipt and document preparation

When completed questionnaires reached the Regional Processing Centres, they were logged, counted and prepared for key entry. Preparation included consistency checks between the questionnaires and the Visitation Record - making sure, for example, that the number of household members on both documents matched. Legibility checks ensured that the documents were suitable for computer entry. Finally, all written answers on household relationships (Question 2) were converted to numerical codes.

(b) Reverse Record Check

A sample of persons was selected from the 1986 Census records and external sources, and 1991 documents were searched for these same persons. If a person was found, 1991 characteristics were noted and sent to head office. For those not found, further tracing determined if they had been enumerated elsewhere in Canada or missed altogether. The results of these searches are coded and captured and the file is turned over to the Social Survey Methods Division for weighting and production of undercoverage estimates.

(c) Economic Coding

Written responses for some labour market questions on the long census forms were converted into numeric codes suitable for direct data entry. Three tasks were involved:

- editing to determine if the respondent had worked at any time during the period of January 1, 1990 to June 4, 1991;
- converting the industry, occupation and place of work to numeric codes;
- editing the class of worker question.

Supervisors and coding consultants resolved any discrepancies in coding before the questionnaires for an enumeration area (EA) proceeded to the next stage. Sometimes other sources, city directories and subject-matter personnel for example, were consulted.

(d) Processing

Questionnaires were transferred in work units for direct data entry at Revenue Canada – Taxation regional processing centres: from there, after keying, they were sent to Statistics Canada in Ottawa.

Direct Data Entry (DDE)

The data entry activity was completed on behalf of Statistics Canada by Revenue Canada – Taxation (RCT). Questionnaire data was key entered at seven (7) RCT regional centres, transmitted to RCT Headquarters in Ottawa and stored on tape cartridges. This operation employed approximately 1,500 people sworn to secrecy under the *Statistics Act*.

Head Office Processing (HOP)

Head Office Processing is a combination of automated and manual processing designed to carry out structural edits on the census data and to process special enumeration returns. Included are returns for Canadians overseas, temporary residents and merchant and navy ships personnel. HOP also processes coverage study returns such as **Reverse Record Check (RRC)**, **Vacancy Check (VC)**, and **Overcoverage Study (OC)**. In addition, HOP is responsible for the preliminary and final population and dwelling counts and for the microfilming of census questionnaires for archival purposes. This operation employs approximately 150 people and is conducted in the Statistics Canada head office in Ottawa.

Head Office Processing consisted of four (4) major activities performed in three (3) phases:

- **DA I – Receipt, Registration and Storage**

Visitation Records and questionnaires for each enumeration area were received, registered and stored at the head office. Tapes containing respondent data were copied and loaded onto the HOP database.

- **DA II – Data Analysis**

Automated structural edits were carried out at the enumeration area, household and person levels, and inconsistencies, such as person count conflicts and household number conflicts were resolved manually.

- **DA III – Special Processing**

Special enumeration returns from Canadians living outside Canada, temporary residents and persons aboard merchant, naval and coast-guard vessels were adjusted to include them. In addition, coverage study returns for checking vacant dwellings, under- and overcoverage were processed, and adjustments were done to the data based on the results of the vacancy check.

- **DA I and/or DA II**

HOP was also responsible for the preliminary and final population and dwelling counts and for the microfilming of census questionnaires for archival purposes.

Automated Coding (AC)

The automated coding operation converted written responses to questions on mother tongue, home language, knowledge of other languages, registered Indian status, place of birth, ethnic origin, major field of study, religion and place of residence 1 year ago and 5 years ago to numeric codes.

The responses were matched against an automated reference file/classification structure containing a series of words or phrases to obtain corresponding numeric codes. An analysis was conducted for each variable to ensure data quality objectives were maintained prior to transferring the records to edit and imputation.

The reference file provided for different spellings of various languages. For example, whether Japanese was keyed as "Japanese" or as "Japenese" was irrelevant since both were assigned the same code. In this way, spelling errors by the respondent, or keying errors had a minimal effect on the coded results.

The automated coding procedure was supplemented by manual coding whenever the response could not be matched with the reference file. The vast majority of the write-in responses to mother tongue, home language and knowledge of non-official languages questions were coded using the automated procedure. In fact, 94% of the write-ins for home language, 95% of the write-ins for mother tongue, and 96% of the write-in responses for knowledge of non-official languages, were processed using this approach. The remaining write-in responses (4% to 6%, depending on the variable) were coded manually.

Table 4.1 Write-in Responses by Method of Coding for Mother Tongue, Home Language and Knowledge of Non-official Languages, Canada, 1991

Variable	Number of responses coded			Distribution of responses coded		
	Automated coding	Manual coding	Total	Automated coding	Manual coding	Total
				%	%	%
Mother tongue	958,669	51,655	1,010,324	94.9	5.1	100.0
Home language	572,609	38,978	611,587	93.6	6.4	100.0
Knowledge of non-official languages	1,211,570	56,093	1,267,663	95.6	4.4	100.0

Source: 1991 Automated Coding - MIS Executive Summary, Week Ending February 7, 1992

V. Edit and Imputation

In the edit and imputation phase, all remaining errors, discrepancies, inconsistencies and missing answers are identified and corrected (including imputation) by a fully automated series of computer programs. The final set of usable "clean" data (free of invalid, inconsistent and missing responses) is produced, comprising a unique database which provides Canada's most detailed information about the population and its characteristics, ranging from the national to the neighbourhood level.

Errors found at this stage can be the result of respondents answering the questions incorrectly or incompletely, or they can be due to errors generated during coding activities and data capture. After errors are detected, values for missing or incomplete entries are imputed. Imputation, which is the correction of the errors, is done using either a "deterministic" or a "hot deck" method. For deterministic imputation, errors are corrected by inferring the appropriate value from answers to other questions. The "hot deck" approach selects a record that has a number of characteristics in common with the record in error, and imputes the missing information from this "donor" record.

Two (2) automated systems are used for editing and imputing census data:

- **CANEDIT**

This system is used to correct the 100% demographic data for age, sex, relationship and marital status, and the labour data from the 20% sample (Form 2B).

- **SPIDER (System for Processing Instructions from Directly Entered Requirements)**

The SPIDER system was developed for the 1981 Census to handle the more complex coded variables and absolute values such as income. Most of the questions asked of 20% of the population are processed using SPIDER.

Edit and Imputation of Language Variables

The edit and imputation of language variables was designed to correct for missing, incomplete, or incompatible responses. The mother tongue data were edited first, followed by data on home language, knowledge of official languages and knowledge of non-official languages. The description of edit and imputation procedures for mother tongue can be found in Catalogue No. 92-335E. The data on mother tongue appear in this section for comparison purposes.

Edit and Imputation of Home Language

The edit and imputation procedure for home language was designed to correct erroneous responses and non-response. In general, the approach for edit and imputation of home language was similar to the method used for mother tongue. The following conditions were corrected by the edit and imputation system:

1. **Non-Response**

This was the most frequently occurring condition corrected during edit and imputation. In total, 315,285 records were imputed for non-response.

These records were imputed from the language characteristics of other members of the household in many cases. If, for example, one family member did not provide a response for home language, and all other family members indicated French as home language, the non-response was imputed from other family members.

Example : Family Relationship	Home Language	
	Before Imputation	After Imputation
Father	French	French
Mother	French	French
Child	Blank	French

Of the records imputed for non-response, 137,620 were imputed using "family imputation", that is, their home language was imputed from that of other family members.

If the respondent was not in a census family or if the home languages reported by other family members were not identical, then the "hot-deck" method of imputation was employed. This procedure involved a search for a "donor" record from which the home language could be imputed to correct for non-response.

The search for a donor record attempted to match the characteristics of the donor and those of the record to be imputed. Records were first stratified into an aboriginal or non-aboriginal stratum. The system then attempted to match on the mother tongue of the record to be imputed. If an identical mother tongue was found, the procedure then tried to find a match based on age and family status. However, the same age and family status characteristics were not considered "mandatory" for the donor record. Only the match on mother tongue was mandatory. A total of 169,255 records were imputed by the hot-deck method, using donor records.

If a suitable donor, with an identical mother tongue, could not be found in the hot-deck, then a default procedure was used. This procedure simply assigned the response appearing as mother tongue to the home language variable. Only 8,410 records were processed in this manner.

2. Same Code Provided for Two Written Responses

If two written responses were coded to the same code for a particular respondent, then one was removed.

3. A Pseudo-code Was Assigned

The pseudo-codes of "none", "Canadian" and "Uncodable" were dealt with in the same way as non-response. In total, 29,820 records were identified with one of these three pseudo-codes.

The pseudo-code of "baby" occurred much more frequently on the database (58,015 records). In these cases, the procedure imputed the same value as the mother tongue response.

The pseudo-code of "Indian" appeared only 40 times on the database. These responses, when reported alone as single responses, were assigned to Aboriginal or Indo-Iranian categories, depending on the value for mother tongue.

4. Write-in Corresponding to English and/or French

When English, French or both were coded as a write-in, the edit and imputation procedure would ensure that the response was transferred to the appropriate mark-in value. This correction was applied to 83,025 records.

5. A Pseudo-language Was Reported

Pseudo-languages refer to write-in responses of either Belgian, Czechoslovakian, Scandinavian or Swiss (not languages per se). In some cases, the resolution mechanism assigned the mother tongue to the home language variable. For example, if the home language was Belgian and the mother tongue was Flemish, then Flemish was assigned as home language. In other cases, a probabilistic algorithm was employed to assign a language. For example, the pseudo-language "Belgian" was assigned to French, Flemish or Dutch based on pre-set probabilities. In total, 2,255 pseudo-languages were corrected.

These five error conditions covered the vast majority of the edit and imputation corrections undertaken for the home language variable. Table 5.1 provides a summary of the edit and imputation error identification for both mother tongue and home language.

Table 5.1 Corrections of Anomalies and Non-response, Mother Tongue and Home Language, Canada, 1991

	Mother tongue		Home language	
	Number	%	Number	%
Total	26,994,045	100.0	26,994,045	100.0
Pseudo-languages	11,555	0.0	2,255	0.0
Uncodable responses	5,030	0.0	3,390	0.0
Responses of "None"	23,010	0.0	25,290	0.0
Responses of "Baby"	27,305	0.0	58,015	0.2
Responses of "Canadian"	1,185	0.0	1,140	0.0
Responses of "Indian"	180	0.0	40	0.0
Write-in official languages	79,930	0.3	83,025	0.3
Family imputation	178,705	0.7	137,620	0.5
Hot-deck imputation	155,015	0.6	169,255	0.6
Hot-deck default deck	155	0.0	8,410	0.0
Other home language correction	N/A		550	0.0
No imputation	26,560,550	98.4	26,530,420	98.3

Note: Based on weighted 20% sample data. Some responses are counted twice.

Edit and Imputation of Knowledge of Official Languages

The edit and imputation procedure for the knowledge of official languages variable was designed to correct for non-response, multiple response and responses that were inconsistent with those for other language variables.

1. Non-response

Non-response was the most frequently occurring error condition corrected for this variable. In total, 191,630 records were imputed for non-response.

The hot-deck approach was employed for imputation. The donor sought in the hot-deck was required to have the same home language as the record requiring imputation. This was the mandatory requirement for imputation using the hot-deck method. A match on age and family status was also sought, but having a donor record with characteristics corresponding to these was not compulsory. In most cases (185,720), a donor record was found, but for a relatively small number of records (5,910), the default imputation was used and values were assigned based on pre-set probabilities.

2. Correction of Inconsistencies

2.1 Inconsistency Between Mother Tongue, Home Language and Knowledge of Official Languages

The correction for inconsistency among these three variables was based on the principle that the respondent must be able to conduct a conversation in a language reported as both mother tongue and home language. When the same official language appeared as both mother tongue and home language, and knowledge of only the other official language was reported for the knowledge of official languages variable, the response to the latter variable was changed to "English and French". When the same official language was reported as both mother tongue and home language, and the record showed knowledge of neither English nor French, the response was changed to correspond with the responses to mother tongue and home language, i.e. to "English only" or to "French only". In total, 158,695 records were corrected due to inconsistency of these three language variables.

2.2 Inconsistency Between Home Language, Knowledge of Official Languages and Knowledge of Non-official Languages (population 2 years of age and older)

If a record showed that a respondent spoke one or both official languages at home, but indicated knowledge of neither official language, the response to the knowledge variable was changed to correspond to the home language. This change was not implemented when the response to the question on knowledge of non-official languages showed that the respondent could speak one or more languages other than English and French. Only 1,515 records were affected by this modification.

2.3 Inconsistency Between Home Language and Knowledge of Official Languages (population less than 2 years of age)

For children less than two years of age, there is a tendency to indicate that they speak neither English nor French, even though there are instructions to report the language used in the home (often English or French) for children who have not learned to speak a language. Consequently, a modification was made if the response indicated that English, French or both languages were reported as home language for the child. In total, 7,065 records were affected by this modification.

2.4 Inconsistency Between Home Language, Knowledge of Non-official Languages and Knowledge of Official Languages (all age groups)

This modification was implemented when the response showed a home language involving one or both official languages as well as a non-official language. If there was no non-official language indicated in response to the question on knowledge of non-official languages, and the respondent indicated knowledge of neither English nor French, the latter response was changed to show a response of English, French or both. There were 3,215 records affected by this modification.

3. Resolution of Multiple Responses

The response categories for the question on knowledge of official languages are mutually exclusive – only one response should be provided. Nevertheless, some respondents provide more than one response. For example, some indicate that they speak "English only" and "French only". The resolution procedure calls for a change to "Both English and French" in such cases. In fact, any multiple response which showed knowledge of both official languages was changed to "Both English and French". In cases where responses indicated "English only" or "French only" along with "Neither English nor French", only the former response (knowledge of an official language) was kept. In total, 174,805 multiple responses were resolved during the edit and imputation process.

Table 5.2 shows a summary of the error conditions corrected during the edit and imputation process.

Table 5.2 Correction of Anomalies and Non-response, Knowledge of Official Languages, Canada, 1991

	Number	%
Total	26,994,045	100.0
Inconsistency - M.T., H.L. and K.O.L.	158,695	0.6
Inconsistency - H.L. and K.O.L. (population 2+)	1,515	0.0
Inconsistency - H.L. and K.O.L. (population L.T. 2)	7,065	0.0
Inconsistency - H.L., K.O.L. and K.N.O.L.	3,215	0.0
Resolution of multiples	174,805	0.6
Hot-deck imputation	185,720	0.7
Hot-deck default	5,910	0.0
No imputation	26,457,120	98.0

Notes: 1. M.T. refers to mother tongue.
H.L. refers to home language.
K.O.L. refers to knowledge of official languages.
K.N.O.L. refers to knowledge of non-official languages.

2. Based on weighted 20% sample data.

Edit and Imputation of Knowledge of Non-official Languages

The edit and imputation for the knowledge of non-official languages variable was designed to correct for non-response, erroneous and incompatible response. There was a relatively high non-response rate for this variable (3.8%), and consequently considerable imputation was necessary.

1. Non-response

The hot-deck method of imputation was not employed to correct for non-response to the question on knowledge of non-official languages. Instead, imputation was carried out after considering the other language characteristics of the respondent. For respondents with non-response, the response was assigned from the home language or from the mother tongue variable. If there was no non-official language reported as home language or mother tongue, a value of "none" was assigned for the knowledge of non-official languages variable, indicating that no non-official languages were known. This was the most frequently occurring type of non-response corrected by the edit and imputation procedure. In total, 918,875 records were assigned a value of "none".

If there was a non-official language (or languages) reported as home language, that response was imputed to the knowledge of non-official languages question. If there was not a non-official language reported as home language, then the mother tongue response was verified and any non-official language found was imputed for the knowledge of non-official languages question. In total, 54,035 responses were imputed from home language and 41,395 from mother tongue.

2. Inconsistent Response

The first type of inconsistent response occurs when a respondent has indicated no knowledge of an official language and has reported "none" for knowledge of non-official languages. This is incompatible with the notion that respondents should have knowledge of at least one language. When such a situation occurred, and one (or more) non-official language(s) was reported as home language, the home language response was imputed to the knowledge of non-official languages variable. In total, 25,555 such inconsistencies were corrected.

The second type of inconsistent response occurs when a respondent has indicated "none" for the knowledge of non-official languages variable but has indicated that he or she speaks a non-official language (or languages) most often at home. In such cases, the home language response is assigned to the knowledge of non-official languages variable. In total, 23,915 such inconsistencies were corrected.

3. Resolution of Multiple Responses

A multiple response for the question on knowledge of non-official languages occurred when the response indicated no knowledge of non-official languages (the box "none" was checked), yet a non-official language was reported as a write-in. In such cases, it was assumed that the write-in was correct, and the response indicating no knowledge of languages other than English or French was changed. In total, 50,560 records were corrected using this procedure.

4. Other Anomalies

In cases where a response was uncodable, or where it showed a value of "none", "baby" or "Canadian", the record was treated as a non-response. Approximately 10,000 records fell into one of these categories.

Responses of "Indian" were dealt with by assigning some to aboriginal languages and some to indo-iranian languages after considering the responses to the home language and mother tongue questions.

Write-ins of English or French for the question on knowledge of non-official languages were eliminated. If this was the only response to the question, the record was treated as a non-response. In total, 39,150 write-ins involving the official languages were reported.

Table 5.3 Correction of Anomalies and Non-response, Knowledge of Non-official Languages, Canada, 1991

	Number	%
Inconsistency - Type 1	25,555	0.1
Inconsistency - Type 2	23,915	0.1
Resolution of multiples	50,560	0.2
Pseudo-languages	14,685	0.1
Uncodable responses	2,760	0.0
Responses of "None"	3,140	0.0
Responses of "Baby"	6,640	0.0
Responses of "Canadian"	360	0.0
Responses of "Indian"	140	0.0
Write-in official languages	39,150	0.1
Imputation from home language	54,035	0.2
Imputation from mother tongue	41,395	0.2
Imputation for "None"	918,875	3.4
No imputation	25,812,835	95.6

Note: Based on weighted 20% sample data.

The Impact of Edit and Imputation on Language Variables

The overall impact of the edit and imputation of data can be studied by comparison of the counts before and after this process.

Tables 5.4 to 5.6 show the distributions before and after edit and imputation for home language, knowledge of official languages and knowledge of non-official languages.

1. Impact of Edit and Imputation - Home Language

Table 5.4 shows the distribution for selected home languages before and after imputation, as well as the percentage imputed. The latter figure is calculated by excluding responses that were modified prior to the imputation process, such as those that were changed because the same language was reported twice for the same person.

Table 5.4 Home Language Before and After Edit and Imputation, Canada, 1991

	Before edit and imputation	After edit and imputation	% edited and imputed	% imputed	% change
Home Language					
English	17,928,005	18,220,180	1.6	1.1	1.6
French	6,146,015	6,211,240	1.1	0.8	1.1
Non-official language	2,012,240	2,074,895	3.0	3.0	3.1
English and French	94,830	113,185	16.2	1.5	19.4
English and non-official language	298,050	320,050	6.9	2.1	7.4
French and non-official language	31,595	33,710	6.3	2.8	6.7
English, French and non-official language	10,635	11,225	5.3	2.6	5.5
Non-official languages	9,055	9,560	5.3	4.6	5.6
Ten Largest Non-official Languages					
Chinese	380,695	389,235	2.2	2.0	2.2
Italian	237,530	241,420	1.6	1.4	1.6
Portuguese	130,615	133,815	2.4	2.1	2.4
Spanish	118,930	123,130	3.4	3.2	3.5
German	111,295	114,270	2.6	2.1	2.7
Polish	101,370	103,905	2.4	2.3	2.5
Punjabi	97,405	101,260	3.8	3.4	4.0
Greek	77,885	79,340	1.8	1.7	1.9
Vietnamese	69,140	71,550	3.4	3.2	3.5
Arabic	62,585	64,935	3.6	3.3	3.8
Languages With Imputation Rates of 5% or Over					
Khmer (Cambodian)	11,795	12,425	5.1	5.0	5.3
Frisian	360	375	4.0	5.3	4.2
Sinhalese	900	945	4.8	5.3	5.0
Danish	1,520	1,625	6.5	5.8	6.9
Cree	47,360	50,765	6.7	6.5	7.2
Kurdish	975	1,045	6.7	6.7	7.2
Montagnais-Naskapi	6,490	7,000	7.3	7.2	7.9
Swedish	1,370	1,495	8.4	8.0	9.1
Icelandic	155	170	8.8	8.8	9.7
Swahili	425	480	11.5	9.4	12.9
Welsh	140	150	6.7	10.0	7.1
Malecite	40	50	20.0	10.0	25.0
Tsimshian	45	45	0.0	11.1	0.0
Norwegian	620	705	12.1	12.1	13.7
Ojibway	10,580	12,100	12.6	12.3	14.4
Gaelic	135	160	15.6	18.8	18.5
North Slave (Hare)	15	15	0.0	33.3	0.0
Mohawk	20	30	33.3	33.3	50.0
Haida	10	5	-100.0	100.0	-50.0

- Note:
1. The % edited and imputed shows the percentage of records changed by the edit and imputation process. It includes changes made to correct invalid responses.
 2. The % imputed includes only those imputed by family, hot-deck or default imputation.
 3. The percentages are calculated from rounded figures and consequently, when the absolute numbers are small, the % imputed is higher than the % edited and imputed in some cases.
 4. The % change refers to the percentage change in figures before and after the edit and imputation process.
 5. Based on weighted 20% sample data.

The percentage of records imputed for the two official languages is quite low (1.1% for English and 0.8% for French). The imputation rate was higher for the major non-official languages, ranging from 1.4% for the Italian home language to 3.4% for Punjabi. A number of languages had imputation rates of 5% or over. In most cases, these languages were spoken by a fairly small number of respondents. The main exceptions were Cree (50,765), Khmer (Cambodian) (12,425), Ojibway (12,100) and Montagnais-Naskapi (7,000).

2. Impact of Edit and Imputation - Knowledge of Languages

The basic distributions before and after edit and imputation for the two knowledge of languages variables are shown in Tables 5.5 and 5.6. The process for knowledge of official languages resulted in an increase of 4.4% in the number of respondents speaking both official languages. A sizeable proportion of this increase was attributable to the resolution of multiple responses that were indicative of a bilingual respondent (for example, a response of "English only" and "French only"). The number speaking neither official language declined considerably during the edit and imputation process. Part of the decline is due to responses that were changed as a result of incompatibility with responses to mother tongue and home language questions.

Table 5.5 Knowledge of Official Languages Before and After Edit and Imputation, Canada 1991

Knowledge of Official Languages	Before edit and imputation	After edit and imputation	% imputed	% change
Total	26,630,160	26,994,040	1.3	1.4
English only	17,838,563	18,106,760	1.5	1.5
French only	4,059,991	4,110,305	1.2	1.2
English and French	4,206,791	4,398,655	4.4	4.6
Neither English nor French	524,815	378,320	-38.7	-27.9

Note: Based on weighted 20% sample data.

The largest changes for the knowledge of non-official languages variable occurred for the two categories "no non-official language" and "one non-official language". Much of the growth in the former category resulted from respondents whose answers to the language questions did not give any indication of knowledge of a language other than English or French, and who did not provide a response to this question. They were assigned a value of "none" during the edit and imputation process.

Table 5.6 Knowledge of Non-official Languages Before and After Edit and Imputation, Canada, 1991

Knowledge of Non-official Languages	Before edit and imputation	After edit and imputation	% imputed	% change
Total	25,929,180	26,994,045	3.9	4.1
No non-official language	21,143,030	22,061,905	4.2	4.3
One non-official language	4,207,100	4,349,260	3.3	3.4
Two non-official languages	452,910	456,355	0.8	0.8
Three or more non-official languages	126,140	126,525	0.3	0.3

Note: Based on weighted 20% sample data.

Weighting

One in every five households or 20% of the population receives a more detailed long questionnaire (Form 2B) and is asked additional socio-economic questions. A weighting algorithm is developed so that these data can be used to estimate response from 100% of the population. The procedure to weight sample data in 1991 has been revised from the 1986 and is known as the "Generalized Least Squares Estimation Procedure (GLSEP)". The GLSEP begins with initial weights of approximately 5 and then, using basic census information known for every person, i.e., age, sex and marital status, adjusts them to obtain the desired agreement between the sample estimates and the population counts. Once data are finalized and weights are calculated, final data are transferred to the Canada Retrieval Databases; these database are used to produce the published and custom products.

VI. Data Evaluation

Throughout the census-taking process, care was taken to ensure high-quality results. Rigorous quality standards were set for data collection and processing, and the Public Communications Program assisted in minimizing non-response. A Data Quality Measurement Program was established to provide users with information on the quality and limitations of census data.

Although considerable effort is made throughout the entire process to ensure high standards of data quality, resulting data are subject to a certain degree of inaccuracy. To assess the usefulness of census data for their purposes and to understand the risk involved in drawing conclusions or making decisions on these data, users should be aware of their inaccuracies and appreciate their origin and composition.

Error can arise at virtually any stage of the census process from preparation of materials to data collection and through the various processing stages. Some errors occur at random and tend to cancel each other out when individual responses are aggregated for a large group. For errors of this nature, the larger the group the more accurate the corresponding estimate and therefore it is important to be cautious when dealing with estimates derived using small aggregated groups of responses. On the other hand, some errors occur more systematically and errors are more serious to data users than random errors.

For census data in general, the principal types of errors are as follows:

- **Coverage Errors**

Occur when individuals and/or dwellings are missed, incorrectly included, or double counted;

- **Non Response Errors**

Occur when responses are not available from some households and/or individuals due to extended absence or for other related reasons;

- **Response Errors**

Occur when respondents, or in some instances Census Representatives, misinterpret a census question and record an incorrect response;

- **Processing Errors**

Can occur during **coding**, when write-in responses are transformed into numerical codes by clerks, **data capture**, when responses are transferred from questionnaires to computer tapes by key entry operators, and **imputation**, when a valid, but not necessarily correct, response is inserted by the computer into a record to replace missing or invalid data;

- **Sampling Errors**

Only apply to supplementary questions on the long (2B) questionnaire, asked of only a twenty percent (20%) sample of households, arise due to the fact that they are weighted to represent the whole population and inevitably differ somewhat from results that would have been obtained had the questions been asked of the total population.

All of the above errors have both random and systematic components. Usually the systematic component of sampling errors is very small in relation to its random component. For other non sampling errors both random and systematic components may be significant.

Four (4) studies are undertaken to measure coverage errors:

- Vacancy Check
- Temporary Residents Study
- Reverse Record Check
- Overcoverage Study

Two (2) studies are conducted to evaluate response errors:

- Reverse Record Check Content Study
- Overcoverage Content Study

Four (4) studies are undertaken to evaluate the effect of sampling errors on the sample data:

- Sampling Bias Study
- Weighting Evaluation
- 2A/2B Consistency Study
- Sampling Variance Study

Two further studies are done to evaluate the data:

- Edit Sample Study
- Clustered Non-response Study

Besides these studies, before the data are approved for dissemination, the Certification Task is performed to detect any anomalies not identified during the other stages. This process involves analysing the evolution of trends for the variable to be published.

On some Indian reserves and settlements, (total of 78), enumeration was not permitted, was interrupted before completion or the quality of collected data was considered to be inaccurate. These areas are called **Incompletely Enumerated Indian Reserves and Indian Settlements**. Under these circumstances, data are not available for these areas, not included in tabulations, and are noted accordingly where applicable. Caution should be exercised when analyzing data from areas affected by incomplete enumeration especially in small areas where the impact is the greatest.

The inclusion of **non-permanent residents** in the 1991 census will affect the variables that were collected on a one hundred percent (100%) basis such as age, sex, mother tongue, and marital status. Census data on immigration will have to be examined carefully to determine the extent of the variations caused by enumerating non-permanent residents.

For additional information on non-permanent residents, please refer to Chapter III on Collection and Coverage.

VII. Historical Comparability

To fully utilize census data we must not only analyze the historical trends of the data we are presenting, but the historical changes in what information is desired and how it is collected. In the past, the Census of Canada has undergone continuing change to meet Canadians ever changing needs for timely and accurate information on Canada's statistical profile. This versatile perspective has endured in 1991.

The census questionnaire was completely redesigned for the 1991 Census. The following changes were made since the 1986 Census:

- Twelve (12) questions not asked during the 1986 Census appear on the 1991 Census Questionnaire;
- Of the twelve (12) questions, seven (7) appeared for the first time and five (5) were reinstated from previous censuses;
- Four (4) questions found on the 1986 Census Questionnaire were excluded from the 1991 Questionnaire;
- Two (2) new Census Questionnaires were added in 1991 (Form 2D - Canvasser Questionnaire and Form 3B - Soup Kitchen Questionnaire).

Form 2D was introduced to enumerate remote northern areas and Indian reserves. It contained the same questions as the Form 2B but was set up to be administered in a person to person environment. Form 3B, an experimental pilot questionnaire, consisted of eleven (11) questions. Interviews were conducted on a person to person basis in a sample of soup kitchens in major Canadian cities. This special enumeration procedure resulted in better coverage in major cities.

For the 1991 Census, Statistics Canada implemented a Canada wide Address Register to improve coverage in urban centres with population of fifty thousand (50,000) and over. It is estimated that the coverage increased by over sixty-eight thousand (68,000) people. Please refer to Chapter III on Collection and Coverage for additional information on the Address Register.

For the first time since 1941, both **permanent and non-permanent residents** of Canada were enumerated. A growing segment of Canada's population, non-permanent residents can create a demand for government services such as schooling, language training, health care, and employment programs. Users should be careful when comparing data from 1991 and previous censuses.

Publicity and advertising of the 1991 Census was viewed in a slightly different light. The Sponsorship Program continued to enlist the voluntary support of corporations, associations and government as in 1986 and previous censuses but to compliment this program Statistics Canada implemented a paid publicity campaign aimed at increasing public awareness of the importance of census data and to encourage Canadians to respond accurately to their census questionnaire in a timely fashion. For the 1991 Census, teachers' kits have been introduced to help promote a greater awareness of the availability and uses of census data produced by Statistics Canada.

Comparability of Language Data

1. The Inclusion of Non-permanent Residents in the Census Population

The inclusion of non-permanent residents in the 1991 Census has an impact on the comparability of these data with those of previous censuses, since all enumerations since 1941 have included only permanent residents. The change has an effect on the study of changes in the language situation, particularly for the 1986-1991 period. In 1991, there were 223,410 non-permanent residents enumerated, and their language characteristics were quite different than those of permanent residents (see Table 7.1). Thus, their inclusion in the population universe for the census affects both the numbers and percentages appearing in the various language groups.

Table 7.1 Mother Tongue, Home Language, Knowledge of Official Languages and Knowledge of Non-Official Languages for Permanent and Non-permanent Residents, Canada, 1991

	Total		Permanent		Non-permanent	
	Number	%	Number	%	Number	%
Mother Tongue						
Total	26,994,040	100.0	26,770,630	100.0	223,410	100.0
Single responses	26,663,790	98.8	26,447,375	98.8	216,415	96.9
English	16,169,875	59.9	16,116,825	60.2	53,050	23.7
French	6,502,865	24.1	6,496,235	24.3	6,630	3.0
Non-official language	3,991,050	14.8	3,834,320	14.3	156,730	70.2
Multiple responses	330,250	1.2	323,250	1.2	7,000	3.1
English and French	91,895	0.3	91,595	0.3	300	0.1
English and non-official language	186,820	0.7	182,110	0.7	4,710	2.1
French and non-official language	22,550	0.1	21,485	0.1	1,065	0.5
English, French and non-official language	5,925	0.0	5,745	0.0	180	0.1
Non-official languages	23,060	0.1	22,320	0.1	740	0.3
Home Language						
Total	26,994,045	100.0	26,770,635	100.0	223,410	100.0
Single responses	26,506,310	98.2	26,296,015	98.2	210,295	94.1
English	18,220,175	67.5	18,134,840	67.7	85,335	38.2
French	6,211,235	23.0	6,202,130	23.2	9,105	4.1
Non-official language	2,074,895	7.7	1,959,035	7.3	115,860	51.9
Multiple responses	487,730	1.8	474,615	1.8	13,115	5.9
English and French	113,185	0.4	112,715	0.4	470	0.2
English and non-official language	320,055	1.2	309,845	1.2	10,210	4.6
French and non-official language	33,710	0.1	31,935	0.1	1,775	0.8
English, French and non-official language	11,220	0.0	11,020	0.0	200	0.1
Non-official languages	9,560	0.0	9,110	0.0	450	0.2
Knowledge of Official Languages						
Total	26,994,045	100.0	26,770,635	100.0	223,410	100.0
English only	18,106,760	67.1	17,941,355	67.0	165,405	74.0
French only	4,110,300	15.2	4,098,235	15.3	12,065	5.4
English and French	4,398,655	16.3	4,376,945	16.3	21,710	9.7
Neither English nor French	378,320	1.4	354,100	1.3	24,220	10.8
Knowledge of Non-official Languages						
Total	26,994,045	100.0	26,770,635	100.0	223,410	100.0
No non-official language	22,012,440	81.5	21,965,505	82.1	46,935	21.0
One non-official language	4,429,515	16.4	4,276,505	16.0	153,010	68.5
Two non-official languages	434,895	1.6	415,870	1.6	19,025	8.5
Three non-official languages	117,190	0.4	112,745	0.4	4,445	2.0

Note: Based on weighted 20% sample data.

The enumeration of non-permanent residents had a considerable impact on the counts and percentages with a non-official language as home language, as well as those with a knowledge of non-official languages. For example, the population with a non-official language as (single) home language rose by 861,755 between 1986 and 1991. A sizeable proportion of this increase (18%) was attributable to the inclusion of non-permanent residents as part of the population universe. The 7.7% of the population reporting a non-official language as home language would have been 7.3% without non-permanent residents.

The inclusion of non-permanent residents had little impact on the distribution of the population by knowledge of official languages, but the impact on the variable for knowledge of non-official languages was more significant, since a large proportion of them (79.0%) were able to converse in a non-official language.

The effect of the inclusion of non-permanent on the counts for specific languages varies considerably since some languages are over-represented in this population. Table 7.2 shows the principal languages (the 10 most numerous, based on mother tongue) affected by the inclusion of non-permanent residents. English was the language reported most frequently as mother tongue and home language by the non-permanent residents. Among the non-official languages, Chinese and Spanish grew considerably as a result of the change in the census universe (26,250 and 21,920 respectively). Other languages reported frequently by non-permanent residents in answer to the questions on mother tongue, home language and knowledge of non-official languages include Tagalog, Tamil, Arabic, Persian, Japanese, French and Portuguese.

Table 7.2 Principal Languages Affected by the Inclusion of Non-permanent Residents, Canada, 1991

	Non-permanent	Permanent	Total
Mother Tongue			
English	53,050	16,116,825	16,169,875
Chinese	26,250	472,595	498,845
Spanish	21,920	155,505	177,425
Tagalog (Pilipino)	11,730	87,975	99,705
Tamil	10,575	19,960	30,535
Arabic	9,490	98,260	107,750
Persian (Farsi)	7,870	32,750	40,620
Japanese	6,780	23,180	29,960
French	6,635	6,496,230	6,502,865
Portuguese	5,005	207,085	212,090
Home Language			
English	85,335	18,134,845	18,220,180
Chinese	22,640	366,595	389,235
Spanish	17,805	105,325	123,130
Tagalog (Pilipino)	3,370	44,760	48,130
Tamil	8,955	14,130	23,085
Arabic	6,865	58,070	64,935
Persian (Farsi)	6,745	24,895	31,640
Japanese	5,350	10,265	15,615
French	9,105	6,202,130	6,211,235
Portuguese	3,975	129,840	133,815
Knowledge of non-official languages			
Chinese	28,415	528,890	557,305
Spanish	27,115	375,315	402,430
Tagalog (Pilipino)	14,580	122,395	136,975
Tamil	12,290	25,030	37,320
Arabic	13,330	151,055	164,385
Persian (Farsi)	8,810	40,565	49,375
Japanese	7,290	38,080	45,370
Portuguese	5,785	248,680	254,465

Note: Based on weighted 20% sample data.

Detailed comparisons of census results for the 1986-1991 period, particularly for non-official languages, should take the effect of the inclusion of non-permanent residents into account.

2. The Grouping of Language Questions (2B questionnaire)

A number of other factors influence the comparability of language data for the 1986-1991 period. The most significant change was the grouping of all language questions together on the 2B questionnaire. This was a break from the previous census tradition of asking the mother tongue question separately from questions relating to home language and knowledge of official languages. The census results demonstrated clearly that when mother tongue and home language are included in a series of language questions, there are far fewer multiple responses for both questions.

Table 7.3 shows multiple response for mother tongue and home language in the last three censuses. The number and percentage of respondents providing a multiple response for mother tongue in 1991 was far lower than 1986 and even lower than 1981, when an instruction on the questionnaire directed respondents to provide only one response. A similar phenomenon is observed for home language, where the percentage reporting more than one language declined from 4.6% to 1.8% of the population.

Table 7.3 Multiple Response for Mother Tongue and Home Language, Canada, 1981, 1986, 1991

	1991		1986		1981	
	Number	%	Number	%	Number	%
Mother Tongue						
Total	26,994,045	100.0	25,022,005	100.0	24,083,495	100.0
Single responses	26,663,790	98.8	24,161,495	96.6	23,543,020	97.8
Multiple responses	330,250	1.2	860,515	3.4	540,475	2.2
English and French	91,900	0.3	307,250	1.2	183,775	0.8
English and non-official	186,820	0.7	463,480	1.9	298,290	1.2
French and non-official	22,550	0.1	34,725	0.1	20,320	0.1
English, French and non-official	5,925	0.0	41,475	0.2	25,675	0.1
Non-official languages	23,065	0.1	13,590	0.1	12,415	0.1
Home Language						
Total	26,994,045	100.0	25,022,010	100.0	23,867,290	100.0
Single responses	26,506,310	98.2	23,862,335	95.4	23,331,555	97.8
Multiple responses	487,730	1.8	1,159,670	4.6	535,740	2.2
English and French	113,190	0.4	351,905	1.4	150,375	0.6
English and non-official	320,050	1.2	712,445	2.8	342,480	1.4
French and non-official	33,710	0.1	40,060	0.2	19,065	0.1
English, French and non-official	11,220	0.0	47,755	0.2	20,605	0.1
Non-official languages	9,560	0.0	7,510	0.0	3,215	0.0

Note: Based on weighted 20% sample data.

Much of the decline in multiples for home language was due to a decrease in the tendency to report a non-official language in conjunction with one or both official languages. The reporting of both official languages as mother tongue also declined between the two census years.

It appears that respondents are better able to understand the language questions when they are grouped together. Because of the much lower proportion of multiple responses in 1991, comparison with 1986 results must be undertaken with caution. In most cases, users are better advised to employ the 1981 data and to study the evolution of languages over the decade, rather than using the 1986-1991 comparison. Similarly, users who study the evolution of language groups, using the mother tongue variable, are also generally better advised to use the 1981-1991 comparison.

Other Changes Affecting Data Comparability

1. Question Wording

The question on knowledge of non-official languages was asked for the first time in 1991, and consequently the issue of comparability is irrelevant for this variable. The wording of the question on knowledge of official languages did not change at all between 1986 and 1991. However, there was a change introduced for the home language question. For the 1991 Census, the question on home language read as follows: "What language does this person speak **most often** at home?". The expression "most often" replaced the term "usually" which was used in the 1986 Census. The instruction which appeared in 1986 ("If more than one language, which language do you speak most often?") was suppressed in 1991.

A description of changes affecting the comparability of data on mother tongue can be found in Catalogue No. 92-335E (*Mother Tongue* technical report).

2. Changes in Response Categories on the Questionnaire

While the response categories for the knowledge of official languages question remained unchanged between 1986 and 1991, there was a modification for the home language question. The response categories "Italian", "Chinese" and "German", which appeared on the 1986 Census questionnaire, were not included in 1991. Instead, response categories were provided for English and French, and a space was provided for the write-in response of non-official languages.

3. Changes in the Classification of Languages

Modifications in the classification of languages, which affect intercensal comparability, were made between 1986 and 1991. A detailed description of the language classification for these two census years (as well as 1981) can be found in Appendix C of the Census Dictionary. In 1991, three language families – Sino-Tibetan, Tai and Austro-Asiatic – were added to the classification list.

The following languages, which did not appear in the 1986 classification, were added for 1991:

Turkish, Baluchi, Gujarati, Kurdish, Marathi, Pashto and Lao.

The following languages appeared in the 1986 classification but not in the 1991 classification:

Kaska, Tahltan, Tutchone, and Yellowknife.

These four languages appear in the category "Athapaskan Languages n.i.e." in 1991.

Finally, some residual categories were added to the existing language families (e.g., Germanic languages, n.i.e., Dravidian languages, n.i.e.).

VIII. Products and Services

Consultation on user needs

Greater emphasis was placed on user consultation for the 1991 Census products and services. Over the course of two years, over 3,000 organizations from private and public sectors were approached to solicit their comments for the proposed product and service line.

The primary objective of the project was to consult with current and potential census data users to evaluate the proposed 1991 product and service line. Client feedback obtained in this way was used to assist census personnel in assessing and determining product features, content, prices, etc.

Consultations varied considerably in format and in terms of numbers and client sectors consulted. For example, some smaller consultations, restricted to Regional Reference Centres and Provincial Focal Points tended to be preliminary investigations of newly-developed product types. At the other end of the scale, the Dimensions Series was the subject of a mail survey to 2,500 users and potential users, as well as cross-country focus group discussions. Another mail survey to more than 200 libraries yielded an 80% response rate and provided valuable insights into concerns librarians had with regard to census products. Most other products were presented for consultation to several dozen users from a variety of sectors, either by means of face-to-face interviews or mail-back questionnaires. In many cases, the Regional Reference Centre staff was heavily involved in the organizing of the consultations, conducting the interviews and providing their own feedback.

Between November 1 and November 15, 1990, eight (8) Focus Groups on Census Data Support Information were surveyed for their comments and recommendations regarding the **1991 Census Technical Report Series**. Suggested fundamental changes and improvements to the product helped meet the needs of current and potential users.

Product content determination

While users overwhelmingly endorsed most products and services presented to them, they also provided valuable critiques. Many of the suggestions confirmed the need for changes already planned. In some cases, this feedback provided evidence that there was less demand for a product and therefore no need for its production. Findings from the Task Force on the Census Custom Products Service resulted in a complete restructuring of service to provide better and more timely service. Consultation on the place of work variable was carried out to determine the interest in and level of funding available for coding to the submunicipal level.

Consultation proved to be an essential exercise in developing the shape and content of the census product and service line, and determining market potential and pricing. Furthermore, the public relations aspect cannot be underestimated: consultation enabled members of the public to preview census output and provided assurance that their input makes a difference.

Marketing of products and services

The 1991 Census Marketing Program ensures that potential data users receive the information they need on census products and services in order to make informed decision. It seeks to reach those individuals or enterprises that rely on census data to inform them of the products and services available from the census database and their potential uses and applications. The national headquarters in Ottawa and the regional reference centres across the country work in partnership to ensure that the largest number of people possible are aware of what the census database has to offer.

The Census Marketing Program assumes these tasks by:

- planning and co-ordinating census data releases and publication releases;
- developing a client-oriented approach to the promotion of the census database;

- maintaining relations with sponsors who provided support prior to June 4, 1991;
- sustaining relations with purchasers of 1986 Census data and of similar Statistics Canada products and services;
- providing sales support and training workshops to present users and potential new users of census data;
- integrating the products and services generated by the census with many other products and services available from Statistics Canada.

For each data release, the Census Marketing Program, ensures that the information relative to the release is available to the general public through many outlets, especially the media. Communications with other government departments is achieved through letters to deputy ministers indicating release highlights as well as through briefing sessions and special lecture presentations. In addition, *The Daily* is sent to every Member of Parliament and Senator informing each of the results of every data release. For the first time, in conjunction with census data releases, classroom activities will be made available to teachers across the country. This will promote awareness of the availability and uses of census data and other products and services provided by Statistics Canada. Ten official data releases are scheduled for the period of April 1992 to April 1993.

Demolinguistic Products from the 1991 Census

The demolinguistic products available from the 1991 Census were designed to meet the needs of a wide variety of users. An attempt was made to maintain historical comparability in data presentation, while also providing very useful data for those interested only in the current language situation.

Products and services can be divided into five categories: data products, analytical products, reference products, tabulation services, and workshops and seminars.

Data Products

1. The Nation Series

These publications contain various tables on mother tongue, home language, knowledge of official languages, and knowledge of non-official languages for Canada, the provinces and territories and the census metropolitan areas.

- *Mother Tongue* (Catalogue No. 93-313)
- *Mother Tongue: 20% Sample Data* (Catalogue No. 93-333)
- *Home Language and Mother Tongue* (Catalogue No. 93-317)
- *Knowledge of Languages* (Catalogue No. 93-318)

2. The Dimension Series

This series was composed of two demolinguistic publications presenting tables on first official language for Canada, provinces and territories, census metropolitan areas, census divisions and census subdivisions, and on language retention and transfer for Canada, provinces and territories and the largest census metropolitan areas.

- *Population Estimates by First Official Language Spoken, 1991* (Catalogue No. 94-320)
- *Language Retention and Transfer, 1991* (Catalogue No. 94-319)

3. Basic Summary Tabulations

Basic summary tabulations contain data on mother tongue, home language, knowledge of official languages and knowledge of non-official languages, each shown by age group. There is also a summary tabulation showing mother

tongue cross-classified by home language and sex. These data are available for Canada, the provinces and territories, census metropolitan areas, census agglomerations, federal electoral districts, enumeration areas and census divisions and subdivisions. Data are on magnetic tape.

4. Area Profile Series

This series gives a statistical overview or profile of geographic areas. A wide range of census variables appear in these publications (also available on CD-ROM). Part A of the series contains data on mother tongue, while Part B contains data on home language, knowledge of official languages and knowledge of non-official languages.

Data in this series are available for the following geographic areas: federal electoral districts, census metropolitan areas and census agglomerations, urban and rural areas, census divisions and subdivisions, census tracts, enumeration areas and urban forward sortation areas.

Analytical Products

1. Focus on Canada Series

This series provides analytical, non-technical studies which illustrate the changing social, economic and demographic characteristics of Canadians. The demolingistic publication in this series is:

- *Languages in Canada* (Catalogue No. 96-313)

2. Short Article Series

These articles provide a brief analysis of important trends in Canadian society. One article dealing with a demolingistic topic was produced in this series:

- Non parlo né inglese, né francese (Short Article No. 5)
(I Can't Speak English or French)

3. Census Highlights

This series provides the media with articles of general interest on current social issues. Results of the 1991 Census are illustrated through short articles of a non-technical nature.

Reference Products

The reference products available from the 1991 Census include the *1991 Census Dictionary* (Catalogue No. 93-301E), *1991 Census Catalogue* (Catalogue No. 92-302E), *1991 Census Handbook* (Catalogue No. 92-305E) and the *Census Tabulation Guide*. The Technical Reports, including *Mother Tongue* (Catalogue No. 92-335E), also appear in this series.

Tabulation Services

Tabulation services provide products and services designed to meet user's needs in a more detailed or specific form than is available from standard products. The services offer varying degrees of flexibility in terms of content, geography level and medium.

Workshops and Seminars

Workshops and seminars are offered to provide census data users with relevant information and to assist them in maximizing the value of census data to meet their needs. A workshop on language data will be available.

IX. Conclusion

This document was designed to provide users of demolinguistic data with information on the quality of the data. The report dealt primarily with the home language and the two knowledge of languages variables, including their concepts and definitions, collection and coverage, data processing methods, historical comparability and the availability of demolinguistic products. Data on mother tongue, taken from the 20% sample, were often shown for comparison purposes.

The following summarizes a number of the observations made in the previous chapters:

While net undercoverage (undercoverage minus overcoverage) was estimated at less than 3% for those with English or French as home language, it rose to 5.7% for those with a non-official language as home language. Estimation of net undercoverage rates for the knowledge of official languages variable resulted in rates of 3% or less for those who speak English only, French only, or Both English and French, while the rate for those who speak neither official language was 10.9%.

The non-response rates for the three traditional language variables were low - 1.2% for mother tongue and home language, and 0.7% for knowledge of official languages. The question on knowledge of non-official languages had a non-response rate of 3.8%, largely because many respondents skipped it completely rather than indicating that they had no knowledge of a non-official language, i.e. checking the circle for "None".

Multiple responses for the home language question were recorded for 1.8% of respondents. By comparison, 1.2% of respondents provided a multiple response to the mother tongue question. The tendency to provide multiple responses was lowest in the eastern provinces, somewhat higher in New Brunswick, and markedly higher in all other provinces and territories.

Non-official languages reported for the demolinguistic questions were coded either by an automated coding procedure or manually. The vast majority were coded using the automated procedure - 95% of mother tongue write-in responses, 94% for home language and 96% for knowledge of non-official languages.

Only a small percentage (1.7%) of the responses to the home language question, and a similar percentage of those for mother tongue (1.6%) were imputed. Corrections of non-response and anomalies for the knowledge of official languages variable affected 2% of the records. Over half of the imputation for this question was attributable to the resolution of multiples and the correction of inconsistencies between the mother tongue, home language and knowledge of official language variables. The imputation rate for knowledge of non-official languages (4.4%) was considerably higher than the rate for other demolinguistic variables. Much of the imputation for this variable resulted from respondents not checking the circle indicating no knowledge of non-official languages.

Comparison of 1991 Census data with those of previous censuses is affected by the inclusion, for the first time since 1941, of non-permanent residents. The effect is greatest for non-official languages, since the majority of the non-permanent residents has a non-official language as mother tongue and home language. The two non-official languages affected most (in absolute terms) by the inclusion of non-permanent residents are Chinese and Spanish.

Comparison of the 20% sample data from the 1991 Census with data from 1986 is complicated by a considerable decrease in the tendency to provide multiple responses. The grouping of the language questions together on the 2B questionnaire was largely responsible for the decline (the rate for home language declined from 4.6% to 1.8%). Users studying the evolution of language groups are generally better advised to use the 1981-1991 comparison rather than 1986-1991.

Regional Reference Centres

Statistics Canada's regional reference centres provide a full range of census products and services. Each reference centre is equipped with a library and a sales counter where users can consult or purchase publications, microcomputer diskettes, microfiche, maps and more.

The staff of the regional reference centres provides consultative and research services in addition to providing after-sales service and support, including seminars and workshops on the use of Statistics Canada information.

Each centre has facilities to retrieve information from Statistics Canada's computerized data retrieval systems CANSIM and E-STAT. A telephone inquiry service is also available with toll-free numbers for regional users outside local calling areas. Call, write, fax or visit the nearest regional reference centre for more information.

Atlantic Region

Serving the provinces of Newfoundland and Labrador, Nova Scotia, Prince Edward Island and New Brunswick.

Advisory Services
Statistics Canada
Viking Building, 3rd Floor
Crosbie Road
St. John's, Newfoundland
A1B 3P2

Toll-free service: 1-800-565-7192
Fax number: (709) 772-6433

Advisory Services
Statistics Canada
North American Life Centre
1770 Market Street
Halifax, Nova Scotia
B3J 3M3

Toll-free service: 1-800-565-7192
Local calls: (902) 426-5331
Fax number: (902) 426-9538

Quebec Region

Advisory Services
Statistics Canada
200 René Lévesque Blvd. W.
Guy Favreau Complex
Suite 412, East Tower
Montréal, Quebec
H2Z 1X4

Toll-free service: 1-800-361-2831
Local calls: (514) 283-5725
Fax number: (514) 283-9350

National Capital Region

Statistical Reference Centre (NCR)
Statistics Canada
R.H. Coats Building Lobby
Holland Avenue
Ottawa, Ontario
K1A 0T6

If outside the local calling area, please dial the toll-free number for your region.

Local calls: (613) 951-8116
Fax number: (613) 951-0581

Ontario Region

Advisory Services
Statistics Canada
Arthur Meighen Building, 10th Floor
25 St. Clair Avenue East
Toronto, Ontario
M4T 1M4

Toll-free service: 1-800-263-1136
Local calls: (416) 973-6586
Fax number: (416) 973-7475

Pacific Region

Serving the province of British Columbia and the Yukon Territory.

Advisory Services
Statistics Canada
Sinclair Centre, Suite 300
757 West Hastings Street
Vancouver, British Columbia
V6C 3C9

Toll-free service: 1-800-663-1551
Local calls: (604) 666-3691
Fax number: (604) 666-4863

Prairie Region

Serving the provinces of Manitoba, Saskatchewan, Alberta and the Northwest Territories.

Advisory Services
Statistics Canada
MacDonald Building, Suite 300
344 Edmonton Street
Winnipeg, Manitoba
R3B 3L9

Toll-free service: 1-800-563-7828
Local calls: (204) 983-4020
Fax number: (204) 983-7543

Advisory Services
Statistics Canada
Avord Tower, 9th Floor
2002 Victoria Avenue
Regina, Saskatchewan
S4P 0R7

Toll-free service: 1-800-563-7828
Local calls: (306) 780-5405
Fax number: (306) 780-5403

Advisory Services
Statistics Canada
First Street Plaza, Room 401
138 - 4th Avenue South-East
Calgary, Alberta
T2G 4Z6

Toll-free service: 1-800-563-7828
Local calls: (403) 292-6717
Fax number: (403) 292-4958

Advisory Services
Statistics Canada
Park Square, 8th Floor
10001 Bellamy Hill
Edmonton, Alberta
T5J 3B6

Toll-free service: 1-800-563-7828
Local calls: (403) 495-3027
Fax number: (403) 495-5318

Telecommunications Device for the Hearing Impaired: 1-800-363-7629

Toll Free Order Only Line (Canada and United States): 1-800-267-6677

Depository Libraries

The Statistics Canada Library in Ottawa maintains complete current and historical records of all Statistics Canada publications, both catalogued and non-catalogued. The library staff is available to help users find the required information.

Statistics Canada Library
R.H. Coats Building, 2nd Floor
Tunney's Pasture
Ottawa, Ontario
K1A 0T6
Local calls: 613-951-8219/20
Fax: 1-613-951-0939

The following is a list of full depository libraries that receive all Statistics Canada publications and all other federal government publications.

Canada

Newfoundland

St. John's

Memorial University of Newfoundland
Queen Elizabeth II Library
St. John's, Newfoundland
A1B 3Y1

Prince Edward Island

Charlottetown

Government Services Library
Charlottetown, Prince Edward Island
C1A 3T2

Nova Scotia

Halifax

Dalhousie University
Killam Memorial Library
Halifax, Nova Scotia
B3H 4H8

Wolfville

Acadia University
Vaughan Memorial Library
Wolfville, Nova Scotia
B0P 1X0

New Brunswick

Fredericton

Legislative Library
Fredericton, New Brunswick
E3B 5H1

University of New Brunswick
Harriet Irving Library
Fredericton, New Brunswick
E3B 5H5

Moncton

Université de Moncton
Bibliothèque Champlain
Moncton, New Brunswick
E1A 3E9

Sackville

Mount Allison University
Ralph Pickard Bell Library
Sackville, New Brunswick
E0A 3C0

Quebec

Montréal

Municipal Library of Montréal
Montréal, Quebec
H2L 1L9

Services documentaires multimédia
Montréal, Quebec
H2C 1T1

Concordia University Library
Montréal, Quebec
H3G 1M8

McGill University
McLennan Library
Montréal, Quebec
H3A 1Y1

Université de Montréal
Bibliothèque des sciences humaines
et sociales
Montréal, Quebec
H3C 3T2

Université du Québec à Montréal
Bibliothèque
Montréal, Quebec
H2L 4S6

Québec

National Assembly Library
Québec, Quebec
G1A 1A5

Sherbrooke

Université de Sherbrooke
Bibliothèque générale
Cité universitaire
Sherbrooke, Quebec
J1K 2R1

Sainte-Foy

Université Laval
Bibliothèque générale
Sainte-Foy, Quebec
G1K 7P4

Ontario

Downsview

York University
Scott Library
Downsview, Ontario
M3J 2R6

Guelph

University of Guelph
Library
Guelph, Ontario
N1G 2W1

Hamilton

Hamilton Public Library
Hamilton, Ontario
L8R 3K1

McMaster University
Mills Memorial Library
Hamilton, Ontario
L8S 4L6

Kingston

Queen's University at Kingston
Douglas Library
Kingston, Ontario
K7L 3N6

London

The University of Western Ontario
D.B. Weldon Library
London, Ontario
N6A 3K7

Ottawa

Library of Parliament
Canadian Government Information
Section
Ottawa, Ontario
K1A 0A9

National Library of Canada
Ottawa, Ontario
K1A 0N4

University of Ottawa
Morisset Library
Ottawa, Ontario
K1N 9A5

Sudbury

Laurentian University of Sudbury
Library
Sudbury, Ontario
P3C 2C6

Thunder Bay

Lakehead University
Chancellor Paterson Library
Thunder Bay, Ontario
P7B 5E1

Thunder Bay Public Library
Thunder Bay, Ontario
P7E 1C2

Toronto

Legislative Library
Toronto, Ontario
M5S 1A5

Metropolitan Toronto Reference
Library
Toronto, Ontario
M4W 2G8

University of Toronto
Robarts Library
Toronto, Ontario
M5S 1A5

Waterloo

University of Waterloo
Dana Porter Arts Library
Waterloo, Ontario
N2L 3G1

Windsor

Windsor Public Library
Windsor, Ontario
N9A 4M9

Manitoba**Winnipeg**

Legislative Library
Winnipeg, Manitoba
R3C 0V8

The University of Manitoba
Elizabeth Dafoe Library
Winnipeg, Manitoba
R3T 2N2

Saskatchewan**Regina**

Legislative Library
Regina, Saskatchewan
S4S 0B3

Saskatoon

University of Saskatchewan
The Main Library
Saskatoon, Saskatchewan
S7N 0W0

Alberta**Calgary**

The University of Calgary
MacKimmie Library
Calgary, Alberta
T2N 1N4

Edmonton

Edmonton Public Library
Edmonton, Alberta
T5J 2V4

Legislative Library
Edmonton, Alberta
T5K 2B6

The University of Alberta
Library
Edmonton, Alberta
T6G 2J8

British Columbia**Burnaby**

Simon Fraser University
Library
Burnaby, British Columbia
V5A 1S6

Vancouver

The University of British Columbia
Library
Vancouver, British Columbia
V6T 1Y3

Vancouver Public Library
Vancouver, British Columbia
V6Z 1X5

Victoria

Legislative Library
Victoria, British Columbia
V8V 1X4

University of Victoria
McPherson Library
Victoria, British Columbia
V8W 3H5

Northwest Territories**Yellowknife**

Northwest Territories
Government Library
Yellowknife, Northwest Territories
X0E 1H0

Other Countries**Federal Republic of Germany**

Preussischer Kulturbesitz
Staatsbibliothek
Abt. Amtsdruckchriften U. Tausch
Postfach 1407
1000 Berlin 30
Germany

United Kingdom

The British Library
London, WC1B 3DG
England, United Kingdom

Japan

National Diet Library
Tokyo, Japan

United States of America

Library of Congress
Washington, D.C. 20540
United States of America



ORDER FORM

Statistics Canada

[illegible]

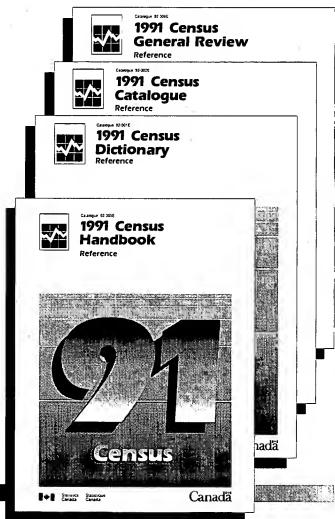
THANK YOU FOR YOUR ORDER!



Statistics Canada / Statistique Canada

Canada

**The Census
is the most
comprehensive
database
available in
Canada today...
are you sure
you're getting
everything
out of it?**



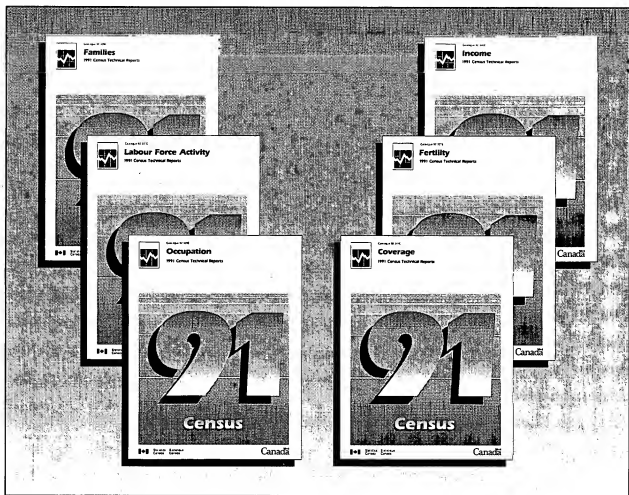
Taking full advantage of Canada's largest, most comprehensive social and economic database is often overwhelming, but the Census can be the most valuable business tool you will ever use. Statistics Canada has designed a series of 1991 General Reference Products to put the Census to work for you.

To order the *Census Dictionary*, *Census Handbook*, *Census General Review* or a *Census Catalogue* of products and services, call your nearest Statistics Canada Regional Reference Centre or our national

☎ 1 800 267-6677 ☎

Nuts and Bolts of Census Data

1991 Census Technical Reports provide users with data quality information. Census concepts, variables and their components, definitions, coverage, processing, data evaluation and limitations and much more are explained in detail in each report.



For a complete list of *Technical Reports* available, call
your nearest Statistics Canada Regional Reference Centre
or our national order line...

 **1 800 267-6677** 